Evaluation of the clinical utility of new diagnostic tests for tuberculosis: the role of pragmatic clinical trials*

Avaliação da utilidade clínica de novos testes diagnósticos em tuberculose: o papel dos ensaios clínicos pragmáticos

Gisele Huf, Afrânio Kritski

Abstract

Tuberculosis is one of the major infectious diseases in developing countries, and the length of time for which the chain of transmission is maintained has been implicated as a major factor in the perpetuation of the disease. In this context, regulatory agencies in such countries have approved new diagnostic tools, which have been almost immediately incorporated into the national tuberculosis control programs. Health interventions have been increasingly investigated in clinical trials, including explanatory trials (in order to evaluate the beneficial effects of such interventions) and pragmatic trials (in order to aid in the decision-making process). We argue that the evaluation of new diagnostic techniques for the detection of tuberculosis should not escape this same logic of evaluation.

Keywords: Tuberculosis/diagnosis; Tuberculosis/prevention & control; Controlled clinical trial as topic.

Resumo

A tuberculose é uma das principais doenças infecciosas em países em desenvolvimento, e o tempo durante o qual a cadeia de transmissibilidade é mantida tem sido apontado como um dos principais fatores de perpetuação da doença. Nesse contexto, novas ferramentas para o diagnóstico têm sido aprovadas pelas agências regulamentadoras nesses países, e sua incorporação é quase imediatamente implementada pelos programas governamentais de controle da tuberculose. As intervenções na área da saúde têm sido cada vez mais investigadas em ensaios clínicos explanatórios (que avaliam seus efeitos benéficos) e em ensaios clínicos pragmáticos (que auxiliam na tomada de decisão). Nós argumentamos que a avaliação de novas técnicas diagnósticas para a detecção da tuberculose não deveria escapar dessa mesma lógica de escrutínio.

Descritores: Tuberculose/diagnóstico; Tuberculose/prevenção e controle; Ensaio clínico controlado como tópico.

General aspects regarding the evaluation of new diagnostic tests for different adverse health events

Although there are various ways and steps to evaluate the potential value of a diagnostic test for clinical use, the choice of study model depends on the question to be answered.

The first question that a new test raises is whether the results obtained in sick individuals differ from those obtained in healthy individuals. In order to answer that question, a study should investigate individuals who are known to have a given disease and those who are healthy, as well as analyzing the results obtained in each group. Such studies evaluate the sensitivity, specificity, and accuracy of the new diagnostic test for a given disease and do not constitute a diagnostic measure; rather, they constitute an initial phase that might contribute to a deeper understanding of the mechanism of action of the disease, as well as aiding in controlling the disease. In that

phase, there is usually greater participation of researchers who are involved in basic and applied basic research, conducted at university research laboratories, research institutes, and industries.[1]

The next question is whether the test can, in suspected cases of the disease in question, distinguish between individuals who have the disease and those who are healthy. In that phase, the evaluation traditionally occurs at clinical research centers and consists of comparing the new test with a reference test, i.e., the gold standard, in order to obtain measurements of diagnostic accuracy, such as sensitivity and specificity. Such studies constitute the vast majority of those evaluating new diagnostic tests and published in recent years, having been conducted at clinical research centers, universities, and research institutes with the support of the industry and (usually) the structural conditions needed in order to meet the demands of regulatory agencies, such as the Brazilian National Health Oversight Agency and the US Food and Drug Administration (FDA).

In order to make health care decisions, including those regarding the diagnostic method to be used, we have to consider the available scientific evidence regarding the risks and benefits of alternative strategies. However, we should also take into consideration the reliability of such evidence. That concern provided the spark for the emergence of a series of formal systems designed to grade the quality of evidence, which can range from very high to very low. Among such systems, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system, adopted by the World Health Organization (WHO), is noteworthy. The GRADE system was initially proposed in Canada with the objective of evaluating new recommendations for tuberculosis control, having come to be more widely used in developed countries in the last decade.[2-4]

The GRADE system has two principal functions. First, it classifies the level of scientific evidence on the basis of the ability of a study to avoid systematic errors or biases, as follows: level 1—clinical trials and systematic reviews/meta-analyses of clinical trials; level 2—observational cohort and case-control studies; level 3—descriptive studies, whether analytical or not, with no comparison group; and level 4—case studies or expert opinions. Second, it grades the strength of the recommendations, meaning that studies with a high level of scientific evidence do not

necessarily entail a high grade of recommendation, and vice-versa. For instance, a disturbing adverse effect could make a given treatment option inappropriate for certain patients, despite the fact that the evidence for that treatment option originated from high-quality randomized studies; in contrast, a given adverse effect documented by observational studies could refrain physicians from prescribing a certain treatment when there are other, equally effective (and safer), options available.

Although the recommendations for the use of diagnostic tests and those for the use of therapeutic interventions follow the same rationale, such tests have peculiar characteristics and pose unique challenges, as evidenced by the application of the GRADE system to the area of diagnostic tests and strategies. The evaluation of the results of new diagnostic tests shows that a low accuracy greatly limits the clinical value of a given test; however, a high sensitivity or a high specificity (or a combination of the two) is not enough to guarantee an improvement in patient-centered and physician-centered outcomes. In practice, clinicians wish to know how much a given test can influence the clinical judgment when deciding on how to treat a given patient. The question is whether the individuals who are tested do better than do those who are not in terms of the outcomes of the diagnostic and therapeutic interventions resulting from the new test. That benefit is rarely evident in the literature; in general, as is the case of tests for early detection of an asymptomatic disease, this can be accurately evaluated only by following individuals who were randomly selected to undergo the test of interest, a different test, or no test.

The use of clinical trials in order to evaluate diagnostic procedures is still in its infancy. In most cases, the best diagnostic strategy is unknown, given that each has advantages and disadvantages; however, the question is whether the new diagnostic strategy should really replace the current strategy, and studies should evaluate the effects that the new strategies have on the decision-making process, as well as on patient-centered outcomes and on the costs to society.

One of the difficulties is that diagnostic tests have no direct impact on patient-centered outcomes; rather, they affect subsequent decisions. The relationship between a given diagnostic test and patient-centered outcomes is predominantly

indirect. The primary desired effect on patient health is rarely produced by the diagnostic test itself. The real interest lies in how the information obtained by the diagnostic test can improve health, meaning that the desired effect is achieved through the mediating role that the diagnostic test plays in the decisions regarding other interventions. The information obtained by a diagnostic test is more commonly used in order to support decisions regarding treatment initiation, modification, or discontinuation. Therefore, a comparison of diagnostic strategies is actually a comparison of strategies that include the diagnostic method and the treatment that it entails. In this case, the value of the diagnostic test is evaluated, as is the benefit of early detection and appropriate treatment of the disease. Therefore, the advantage of randomized studies is that they reduce the bias created by external factors when comparing groups. However, it is fundamental that the results be generalizable; how different are our patients from those included in the studies? Can the results obtained in those studies be applied to our patients?

Only after more than two decades of systematic reviews and meta-analyses was it realized that the outcomes assessed in most clinical trials do not answer the key questions involved in the decision of whether or not to incorporate the technology into the health care system; such studies involve specific populations that are unrepresentative of the general population and, in general, do not include cost-effectiveness evaluations.[5] In the late 1990s, the distinction between explanatory and pragmatic clinical trials began to gain prominence.[6] Explanatory clinical trials, conducted at clinical research centers, seek to answer questions of effectiveness, i.e., whether and how an intervention works, whereas pragmatic trials are conducted to support decision making in the health care area and, to that end, are carried out under conditions that are very close to those under which health care facilities routinely work. Pragmatic trials also involve patients who are very similar to those who will need treatment in the future. Pragmatic trials have received little or no attention from the academic community, government funding agencies, and industries, the last of the three being particularly interested in registering the product with regulatory agencies and marketing it. A recent systematic review of 168,000 randomized

clinical trials conducted between 1976 and 2002 showed that only 95 of the trials (0.05%) met the classification criteria for pragmatic clinical trials. The authors of the review emphasized the urgent need for prioritizing pragmatic trials aimed at answering whether the new technology can be applied to the health care system over explanatory trials focusing exclusively on efficacy, the primary objective of which is to register the product with regulatory agencies.[7]

The situation is not different in the diagnostic testing area. Industry-funded diagnostic accuracy studies conducted at clinical research centers have been enough for new diagnostic tests to be approved by regulatory agencies (the US FDA, the European Agency for the Evaluation of Medicinal Products, and the Brazilian National Health Oversight Agency) for marketing. The new diagnostic tests are incorporated into the private health care system as soon as they are commercially available, on the basis of studies involving a limited number of cases, meaning that the expectations regarding the utility of those tests are subjective. In the current economic system, lobbying by the industry and civil society (the latter being influenced by marketing strategies) has created a situation in which individuals seeking health care are offered all technological innovations for which there is any scientific evidence, however minimal, even if there has been no systematic evaluation of the impact that such innovations have on the health care system.

Evaluation and incorporation of new technologies for the diagnosis of tuberculosis in Brazil and in the world

Tuberculosis is one of the major infectious diseases in developing countries, as well as being a major cause of morbidity and mortality. Tuberculosis transmission has been reported to be a major factor in the perpetuation of the disease worldwide, being associated with factors such as social inequality, the advent of AIDS, and the aging of the population.[8]

The WHO estimated that the overall incidence of tuberculosis in 2009 was 9.4 million cases, the incidence of tuberculosis in that year having therefore increased in comparison with the incidence of the disease in 2000 and in 1990 (8.3 and 6.6 million cases, respectively).[8] Although

the total number of cases of tuberculosis has increased (in absolute terms) with the growth of the population, the number of cases per capita has decreased. However, the rate of decline remains low, at < 1% per year.[8] In 2009, there were approximately 440,000 new cases of multidrug-resistant tuberculosis (MDR-TB), 150,000 of those patients having died. In addition, there were 1,400,000 deaths among HIV-negative patients and 380,000 deaths among HIV-positive patients, which translated to approximately 4,700 deaths per day.

One of the major factors for the emergence of new cases is the length of time for which patients with pulmonary tuberculosis remain without diagnosis and treatment, therefore maintaining the chain of transmission.[9] The diagnostic yield of sputum smear microscopy for AFB is low (40-60%), principally in patients with low bacterial load in respiratory samples, as occurs in HIV-positive patients and in those with other immunosuppressive diseases (30%).[10] For thirty years, the WHO no longer considered tuberculosis research a priority; however, in 2006, through the Stop TB Partnership Second Global Plan to Stop TB, the WHO again recommended that tuberculosis research be conducted in the areas of development and evaluation of new diagnostic methods, drugs, vaccines, and management strategies.[11] An increase in the diagnosis of paucibacillary tuberculosis, tuberculosis/HIV co-infection, and MDR-TB was therefore essential to the success of the plan, given that rapid diagnosis of tuberculosis allows the initiation of pharmacological treatment and reduces the length of time for which the chain of transmission is maintained, therefore reducing the number of individuals infected by those with the disease.[12]

Although the sensitivity of AFB-positive sputum smears is low, sputum smear microscopy for AFB is still one of the tests that are most widely used in order to diagnose drug-susceptible pulmonary tuberculosis in developing countries.[8] However, in most of those countries, culture for mycobacteria, the sensitivity of which is higher (80-85%) than is that of sputum smear microscopy for the diagnosis of drug-susceptible pulmonary tuberculosis, is performed on Löwenstein-Jensen solid medium and is indicated only in selected clinical cases, including cases of treatment failure, persistently negative sputum smear results, and extrapulmonary forms. The major problem with

culture on Löwenstein-Jensen medium is the long incubation period (4-6 weeks), drug susceptibility testing being performed with the culture rather than with the clinical specimen, several more weeks being therefore required in order to obtain the results.[13]

In HIV-positive patients and in children, the WHO recommends that priority be given to the evaluation of patients with respiratory symptoms (cough for more than 2-3 weeks) in order to search for cases of pulmonary tuberculosis, a strategy that has proved inappropriate. Recently, one group of authors identified 267 cases of tuberculosis among 1,748 HIV-positive patients suspected of having tuberculosis and reported that cough for 2-3 weeks constituted a finding that did not aid in the diagnosis. The presence of cough and fever (regardless of the duration), as well as of sweating for more than 3 weeks, was found to have a sensitivity of 93% and a specificity of 36%.[14] In a review article, one group of authors proposed the development and evaluation of clinical and clinical-radiological scores for the diagnosis of pulmonary tuberculosis in children and adults in various epidemiological settings.[15] This underscores the urgent need for evaluating new diagnostic approaches that might have greater impact on the regions that are most affected by tuberculosis/HIV co-infection and MDR-TB.[16] However, when evaluating new diagnostic technologies, we should consider detection strategies that include the analysis of factors associated with access to health care (patient delay, health care system delay, or both).[17]

The WHO has proposed tests that increase the sensitivity of sputum smear microscopy. However, in practice, until 2007, with the universal use of sputum smear microscopy alone, 30-40% of the patients treated at health care facilities in developing countries were treated for tuberculosis without bacteriological confirmation. Also in 2007, in order to respond more effectively to the worldwide emergence of tuberculosis/HIV co-infection and MDR-TB, the WHO recommended the use of liquid culture for the detection of *Mycobacterium tuberculosis* and for drug susceptibility testing, on the basis of a review of the available scientific evidence and of consultation with an expert panel.[18-20] In 2008, the WHO recommended the use of molecular tests for rapid screening of patients suspected of having MDR-TB. That recommendation was

based on systematic reviews, expert opinions, and preliminary results of effectiveness obtained in demonstration projects (phase III/IV) carried out at clinical research centers. Such tests should be reserved for respiratory samples with positive smear microscopy or culture for mycobacteria.[21,22]

A systematic review of studies evaluating new diagnostic tests for tuberculosis demonstrated the lack of methodological rigor of most of the studies.[23] The authors of the review reported that biased results from poorly designed studies could lead to early adoption of new diagnostic tests, with little or no benefit.[23]

In recent years, guidelines for the standardization of study models in the area of infectious diseases have been published; according to those guidelines, in addition to evaluating the accuracy of a new diagnostic test, studies should evaluate various algorithms (i.e., not only those for individual tests), as well as evaluating the relative contributions of the new test to the health care system, the incremental value of the test, the impact of the test on clinical practice (i.e., its impact on decisions regarding treatment), the cost-effectiveness of the test under routine conditions, and the impact that the results obtained with the new test have on patients and society.[24-26]

In 2008, one group of authors analyzed data from the literature and concluded that the new recommendations included in the guidelines for tuberculosis in developed countries—based on the highest level of scientific evidence (GRADE)—were rapidly incorporated into the clinical guidelines adopted in developing countries, few changes being made in order to adjust the guidelines to the health care needs in those countries[27]; in addition, another group of authors reported that most of the clinical guidelines developed in developing countries were not certified by the Advancing Development Guideline, Reporting and Evaluation in Health Care (AGREE).[28] In 2009, the Third Brazilian Thoracic Association Guidelines on Tuberculosis were published, including 24 recommendations based on the level of scientific evidence available. In the diagnostic field, despite the lack of randomized clinical trials confirming the cost-effectiveness and clinical impact of new tests, the guidelines followed the recommendations by the WHO and the FDA, namely the use of liquid culture in nonradiometric automated systems and the

use of commercially available molecular tests, respectively, in paucibacillary respiratory samples for the diagnostic investigation of pulmonary tuberculosis.[29]

A recent survey of 16 countries with a high burden of tuberculosis found that 7 of the countries investigated had adopted the new diagnostic tools recommended by the WHO in 2007.[30] Curiously, none of the 7 countries evaluated the impact of the incorporation of the new technologies, which is in disagreement with what has been proposed by the WHO and the Stop TB Initiative.[31]

Regarding new diagnostic tests for tuberculosis and MDR-TB, liquid culture methods—such as Bactec 960 (Bactec)—and molecular tests—such as EMTD (Gen-Probe), Amplicor (Roche), TB Test (Biometrix), MTBDRplus (Hain Life Sciences), and, more recently, GeneXpert (Cepheid)—have been evaluated and recommended by the FDA in the USA and the corresponding agencies in the European Union, being marketed in those countries. Those tests have also been marketed in the private sector in upper-middle-income countries, such as Brazil. Although there are no reports in the literature regarding pragmatic clinical trials and the cost-effectiveness of those tests for the diagnostic approach to tuberculosis or MDR-TB in developing countries, the Xpert™ MTB/RIF test was recommended by the WHO in December of 2010.[32] The Xpert™ MTB/RIF is a fully automated molecular test with an integrated processing model designed to purify, concentrate, amplify, and identify target sequences of the *rpoB* gene for the diagnosis of resistance to rifampin. The test is applied to sputum samples and provides results after 120 min, without the need for an expert in molecular biology. The results obtained in demonstration studies (phase III) confirmed that the test is highly specific for the diagnosis of tuberculosis and rifampin-resistant tuberculosis. The sensitivity of the test was found to be 72% in sputum samples from patients with negative sputum smear microscopy, similar to that of other molecular tests, such as Amplicor and EMTD. However, the Xpert™ MTB/RIF test can be decentralized to secondary health care facilities, given that it does not need to be performed in a molecular biology laboratory.[33,34]

In recent years, it has become a consensus among policymakers worldwide that emerging countries, such as Brazil, should lead the way

in health technology assessment, as well as in the analysis of the impact of the incorporation of new technologies, focusing on how the new technologies can improve or maintain health. Diagnostic tests should not escape this same logic.

In recent years, Brazil has been the only emerging country in which health technology assessment has been prioritized, the Brazilian National Ministry of Health having created a special committee for the incorporation of technology, a committee that established norms for the sector. [35] New technologies will be incorporated into the public or private health care system only if the studies investigating those technologies are conducted under field conditions in different regions of the country; if they employ the most appropriate study design; if they include an analysis of the impact that the new technologies will have on the health care system; and if they can provide data to support a health policy.

In 2009, the Federal University of Rio de Janeiro School of Medicine, the University of São Paulo School of Medicine, the Federal University of Minas Gerais School of Medicine, the Federal University of Rio Grande do Sul School of Medicine, the Oswaldo Cruz Foundation Institute of Scientific and Technological Communication and Information in Health, and the Brazilian National Ministry of Health Department of Health Care established a partnership, designated PROQUALIS, in order to disseminate information and technology among health care facilities and workers and provide support for issues related to the quality of health care through a collaborative network of public and private universities, together with the Brazilian National Ministry of Health, the Brazilian National Ministry of Education, and the Brazilian National Ministry of Science and Technology.[36] In addition, the PROQUALIS supports clinical practice by providing guidelines and technical/scientific literature that is current and relevant to health professionals in Brazil. The information is selected by a national network of experts. Surprisingly, as the discussions and information exchanges began, the following facts were noticed: the technicians who participate in health technology assessment do not interact with those who devise or modify norms, manuals, and clinical guidelines (the exception being the technicians working in England, Spain, Canada, the Netherlands, and New Zealand); most of the directors of professional associations (medical or

otherwise) in Brazil are unaware of the criteria used internationally in order to grade the quality of guidelines (AGREE) and have different views regarding the development and monitoring of changes in clinical guidelines.

Because any change in clinical guidelines or manuals of standards entails the incorporation of a technology that might be uncritical and harmful to patients, the health care system, or both, it has become a priority to bring researchers working in health technology assessment closer to the professionals who devise norms, manuals, and guidelines, especially those in the field of medical education, with the objective of implementing curricular changes in the medium and long term.

In 2009, researchers affiliated with the *Rede Brasileira de Pesquisa em Tuberculose* (Rede TB, Brazilian Tuberculosis Research Network) played a relevant role in the development of the new version of the Brazilian National Ministry of Health Manual of Epidemiological Surveillance for Tuberculosis. The Rede TB proposed that any change in the manual be considered a technological incorporation, and the proposal was accepted. Therefore, any change in the manual requires an evaluation of its impact on the Brazilian national health care system. It was considered strategic to bring the professionals who develop guidelines closer to those who evaluate health technologies. In the 2009-2010 period, technicians in the Brazilian National Ministry of Health Department of Science and Technology, technicians in the Brazilian National Ministry of Health Department of Health Surveillance, researchers affiliated with the Professor Hélio Fraga Referral Center, researchers affiliated with the Oswaldo Cruz Foundation Clinical Research Institute, researchers under the Federal University of Rio de Janeiro Academic Tuberculosis Program, researchers under the Rio de Janeiro State University Academic Tuberculosis Program, and researchers affiliated with the Rede TB, as well as representatives of the Melinda Gates Foundation, the International Union Against Tuberculosis and Lung Disease, and the United States Agency for International Development, worked together in numerous multidisciplinary activities. Those activities led to the development of two nationwide projects aimed at evaluating the incorporation of new diagnostic tests for tuberculosis and MDR-TB currently marketed in Brazil and available via the private health care system (GeneXpert and

MTBDRplus, respectively). Those operational research projects prioritized pragmatic randomized clinical trials and evaluated the costs of the new tests to patients and the health care system, as well as analyzing scale-up. In 2010, the Brazilian National Ministry of Health provided financial support, and, in 2011, international financial transfers occurred. The primary objective of the platform agreed upon among the participating institutions is to meet the demands of the Brazilian Ministry of Health Committee for the Incorporation of New Technologies in order to identify settings in which those diagnostic tests can be incorporated into the public health care system.

In parallel, in order to aid in the discussion of this issue, the International Union Against Tuberculosis and Lung Disease and its regional partners, such as the Rede TB, have recently published a new proposal for a platform to evaluate the impact of new technologies for the diagnosis of tuberculosis, a platform that will be adopted by the WHO in 2011.[37,38]

In addition to prioritizing operational research through pragmatic randomized clinical trials, cost analysis, and scale-up, the new health technology assessment platform includes aspects related to equity and access, as well as qualitative studies involving users, health professionals, managers, and local/international industry representatives who will be able to identify factors hindering or facilitating the incorporation of new technologies into the various health care systems. The expectation is that the new platform will aid in answering at least one simple question regarding the incorporation and dissemination of new diagnostic methods for tuberculosis: will the new technology be better for the patients and the current health care system in the country? In addition, new techniques are increasingly compared with current techniques in terms of the amount of resources required. Health is no longer the only criterion on which health care decisions are based; social costs, personal displeasure, and the time spent by health care system users, health professionals, or both should also be taken into account.

As previously mentioned, new diagnostic tools are often incorporated into the routine of health care facilities as soon as they are approved for marketing by regulatory agencies—on the basis of their sensitivity and specificity, as well as of ROC curves—in a frenzy that, although often

unjustified from a patient-centered standpoint, certainly appeals to the desire that clinical researchers have for innovation and the desire that manufacturers and laboratories have for marketing the new tools.

One of the criticisms of pragmatic clinical trials is that they prioritize the applicability and generalizability of the results over the internal validity of the study. However, it is inevitable that introducing the perspective of patients in evaluations makes the outcome measures of the study more subjective and therefore more prone to biases. Nevertheless, we must recognize that, regardless of what is being evaluated, the results will always be colored by our perceptions; our observations probably speak more to how we define a problem and categorize the possible outcomes than does any real phenomenon that we might observe. The consequence of this pragmatic approach is that it is impossible to conduct such an investigation in a "purely experimental" environment, if that is ever desirable. Applied basic research, clinical research, operational research, and clinical practice become intertwined, and the primary outcome measures should always be patient health and greater health care system effectiveness in the areas into which a new technology will be incorporated.

References

1. Sackett DL, Haynes RB. The architecture of diagnostic research. BMJ. 2002;324(7336):539-41. PMid:11872558. PMCid:1122451. http://dx.doi.org/10.1136/bmj.324.7336.539

2. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490. PMid:15205295. PMCid:428525. http://dx.doi.org/10.1136/bmj.328.7454.1490

3. Schünemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, et al. An official ATS statement: grading the quality of evidence and strength of recommendations in ATS guidelines and recommendations. Am J Respir Crit Care Med. 2006;174(5):605-14. PMid:16931644. http://dx.doi.org/10.1164/rccm.200602-197ST

4. The periodic health examination. Canadian Task Force on the Periodic Health Examination. Can Med Assoc J. 1979;121(9):1193-254. PMid:115569. PMCid:1704686.

5. Laupacis A, Straus S. Systematic reviews: time to address clinical and policy relevance as well as methodological rigor. Ann Intern Med. 2007;147(4):273-4. PMid:17638716.

6. Coutinho Eda S, Huf G, Bloch KV. Pragmatic clinical trials: an option in the construction of health-related evidence [Article in Portuguese]. Cad Saude Publica. 2003;19(4):1189-93. PMid:12973582.

7. Treweek S, Zwarenstein M. Making trials matter: pragmatic and explanatory trials and the problem of applicability.

Trials. 2009;10:37. PMid:19493350. PMCid:2700087. http://dx.doi.org/10.1186/1745-6215-10-37

8. World Health Organization. Global tuberculosis control: epidemiology, strategy, financing. Geneva: World Health Organization; 2009.

9. Dye C, Williams BG, Espinal MA, Raviglione MC. Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. Science. 2002;295(5562):2042-6. PMid:11896268. http://dx.doi.org/10.1126/science.1063814

10. Sester M, Giehl C, McNerney R, Kampmann B, Walzl G, Cuchí P, et al. Challenges and perspectives for improved management of HIV/Mycobacterium tuberculosis co-infection. Eur Respir J. 2010;36(6):1242-7. PMid:21119204. http://dx.doi.org/10.1183/09031936.00040910

11. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. The Stop TB Strategy - Building on and enhancing DOTS to meet the TB-related Millennium Development Goals. [Adobe Acrobat document, 24p.]. Available from: http://whqlibdoc.who.int/hq/2006/WHO_HTM_STB_2006.368_eng.pdf

12. Sherman LF, Fujiwara PI, Cook SV, Bazerman LB, Frieden TR. Patient and health care system delays in the diagnosis and treatment of tuberculosis. Int J Tuberc Lung Dis. 1999;3(12):1088-95. PMid:10599012.

13. Fair E, Hopewell PC, Pai M. International Standards for Tuberculosis Care: revisiting the cornerstones of tuberculosis care and control. Expert Rev Anti Infect Ther. 2007;5(1):61-5. PMid:17266454. http://dx.doi.org/10.1586/14787210.5.1.61

14. Fair E, Hopewell PC, Pai M. International Standards for Tuberculosis Care: revisiting the cornerstones of tuberculosis care and control. Expert Rev Anti Infect Ther. 2007;5(1):61-5. PMid:17266454. http://dx.doi.org/10.1586/14787210.5.1.61

15. Marais BJ, Raviglione MC, Donald PR, Harries AD, Kritski AL, Graham SM, et al. Scale-up of services and research priorities for diagnosis, management, and control of tuberculosis: a call to action. Lancet. 2010;375(9732):2179-91. http://dx.doi.org/10.1016/S0140-6736(10)60554-5

16. Lawn SD, Ayles H, Egwaga S, Williams B, Mukadi YD, Santos Filho ED, et al. Potential utility of empirical tuberculosis treatment for HIV-infected patients with advanced immunodeficiency in high TB-HIV burden settings. Int J Tuberc Lung Dis. 2011;15(3):287-95. PMid:21333094.

17. Sreeramareddy CT, Panduru KV, Menten J, Van den Ende J. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. BMC Infect Dis. 2009;9:91. PMid:19519917. PMCid:2702369. http://dx.doi.org/10.1186/1471-2334-9-91

18. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. TB diagnostics and laboratory strengthening - WHO policy - The use of liquid medium for culture and DST, 2007. Available from: http://www.who.int/tb/laboratory/policy_liquid_medium_for_culture_dst/en/index.html

19. Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, et al. A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. Health Technol Assess. 2007;11(3):1-196.

20. Cruciani M, Scarparo C, Malena M, Bosco O, Serpelloni G, Mengoli C. Meta-analysis of BACTEC MGIT 960 and BACTEC 460 TB, with or without solid media, for detection of mycobacteria. J Clin Microbiol. 2004;42(5):2321-5. PMid:15131224. PMCid:404614. http://dx.doi.org/10.1128/JCM.42.5.2321-2325.2004

21. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. Tuberculosis (TB) - Moving research findings into new WHO policies. Available from: http://www.who.int/tb/advisory_bodies/research_to_policy/en/index.html

22. World Health Organization. The Global MDR-TB & XDR-TB Response Plan 2007–2008. Geneva: World Health Organization; 2007.

23. Pai M, Ramsay A, O'Brien R. Comprehensive new resource for evidence-based TB diagnosis. Expert Rev Mol Diagn. 2009;9(7):637-9. PMid:19817547. http://dx.doi.org/10.1586/erm.09.48

24. Peeling RW, Smith PG, Bossuyt PM. A guide for diagnostic evaluations. Nat Rev Microbiol. 2006;4(12 Suppl):S2-6. http://dx.doi.org/10.1038/nrmicro1568

25. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Mar 20]. New Technologies for Tuberculosis Control: A framework for their adoption, introduction and implementation. [Adobe Acrobat document, 120p.]; Available from: http://whqlibdoc.who.int/publications/2007/9789241595520_eng.pdf

26. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. Tuberculosis (TB) - TB diagnostics and laboratory strengthening. Available from: http://www.who.int/tb/dots/laboratory/policy/en/index4.html

27. Wei X, Walley JD, Liang X, Liu F, Zhang X, Li R. Adapting a generic tuberculosis control operational guideline and scaling it up in China: a qualitative case study. BMC Public Health. 2008;8:260. PMid:18662410. PMCid:2515317. http://dx.doi.org/10.1186/1471-2458-8-260

28. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. CMAJ. 2010;182(18):E839-42. PMid:20603348. PMCid:3001530. http://dx.doi.org/10.1503/cmaj.090449

29. Conde MB, Melo FA, Marques AM, Cardoso NC, Pinheiro VG, Dalcin Pde T, et al. III Brazilian Thoracic Association Guidelines on tuberculosis. J Bras Pneumol. 2009;35(10):1018-48. PMid:19918635.

30. van Kampen SC, Ramsay AR, Anthony RM, Klatser PR. Retooling national TB control programmes (NTPs) with new diagnostics: the NTP perspective. PLoS One. 2010;5(7):e11649. PMid:20661463. PMCid:2908286. http://dx.doi.org/10.1371/journal.pone.0011649

31. World Health Organization. New Laboratory Diagnostic Tools for Tuberculosis. Geneva: World Health Organization; 2008.

32. World Health Organization [homepage on the Internet]. Geneva: World Health Organization [cited 2008 Jun 16]. Tuberculosis (TB) - Strategic and Technical Advisory Group for Tuberculosis (STAG-TB). Available from: http://www.who.int/tb/advisory_bodies/stag/en/index.html

33. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. N Engl J Med. 2010;363(11):1005-15. PMid:20825313. PMCid:2947799. http://dx.doi.org/10.1056/NEJMoa0907847

34. Van Rie A, Page-Shipp L, Scott L, Sanne I, Stevens W. Xpert(®) MTB/RIF for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype

or hope? Expert Rev Mol Diagn. 2010;10(7):937-46. PMid:20964612. http://dx.doi.org/10.1586/erm.10.67

35. Brasil. Portaria n. 2.587, de 30 de outubro de 2008. Dispõe sobre a Comissão de Incorporação de Tecnologias do Ministério da Saúde e vincula sua gestão à Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Diário Oficial da União, Brasília, nº 212, p 94, 31 de out. de 2008. Seção 1.

36. Proqualis - Aprimorando as práticas de saúde [homepage on the Internet]. Rio de Janeiro: Fundação Oswaldo Cruz, Ministério da Saúde, c2009 [cited 2008 Jun 16]. Available from: www.proqualis.net

37. Mann G, Squire SB, Bissell K, Eliseev P, Du Toit E, Hesseling A, et al. Beyond accuracy: creating a comprehensive evidence base for TB diagnostic tools. Int J Tuberc Lung Dis. 2010;14(12):1518-24. PMid:21144235.

38. Squire SB, Ramsay AR, van den Hof S, Millington KA, Langley I, Bello G, et al. Making innovations accessible to the poor through implementation research. Int J Tuberc Lung Dis. 2011;15(7):862-70. PMid:21682960. http://dx.doi.org/10.5588/ijtld.11.0161

About the authors

Gisele Huf
Researcher. Oswaldo Cruz Foundation National Institute of Quality Control in Health, Rio de Janeiro, Brazil.

Afrânio Kritski
Vice-Director. Federal University of Rio de Janeiro School of Medicine, Rio de Janeiro, Brazil.